DOCUMENT RESUME

ED 482 082 IR 058 768

AUTHOR	Stuer, Peter; Meersman, Robert; De Bruyne, Steven
TITLE	The HyperMuseum Theme Generator System: Ontology-Based Internet Support for the Active Use of Digital Museum Data for Teaching and Presentation.
PUB DATE	2001-00-00
NOTE	15p.; In: Museums and the Web 2001: Selected Papers from an International Conference (5th, Seattle, Washington, March 15-17, 2001); see IR 058 756.
AVAILABLE FROM	Archives & Museum Informatics, 2008 Murray Ave., Suite D, Pittsburgh, PA 15217; e-mail: info@archimuse.com; Web site: http://www.archimuse.com/. For full text: http://www.archimuse.com/mw2001/.
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Exhibits; Foreign Countries; Hypermedia; *Multimedia Materials; *Museums; Secondary Education; World Wide Web
IDENTIFIERS	Europe; Museum Collections; *Virtual Museums

ABSTRACT

Museums have always been, sometimes directly and often indirectly, a key resource of arts and cultural heritage information for the classroom educator. The Web now offers an ideal way of taking this resource beyond the traditional textbook or school visit. While museums around the globe are embracing the web and putting virtual exhibitions, cultural databases and archives on-line, the educator (or user in general) is still facing the daunting task of integrating this material into an active document, course, curriculum or presentation. This paper reports on the construction of a personalized theme creation engine as a possible catalyst to the active use in secondary education in Europe of digital media published on-line by selected museums. The HyperMuseum Theme Generator System (TGS) is part of the HyperMuseum system, a European virtual museum portal. Its function is to assist in the creation of so-called personalized themes. A personalized theme intends to allow the end user to bring together a unique collection of multimedia objects from the HyperMuseum Server, and to create a personalized rendering of the perceived and/or recorded relationships between these objects, realized as a new multimedia document (website, PowerPoint® presentation, hypertext or Word® document etc.). The TGS supports this creative expression both during the discovery phase, exploring the collection and discovering thematic relationships, as well as the realization phase, the construction of the resulting documents. The primary target audience is non-expert users mainly from the secondary education community.This work is partly supported by the Telematics Program (4th Framework) of the European Commission under Project nr. 3088 (HyperMuseum). Includes five figures. (Contains 11 references.) (Author)

Reproductions supplied by EDRS are the best that can be made
from the original document.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

D. Bearman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

PAPERS
Museums and the Web 2001

The HyperMuseum Theme Generator System: Ontology-Based Internet Support For The Active Use Of Digital Museum Data For Teaching And Presentation

Peter Stuer, Robert Meersman, Steven De Bruyne, STARLab VUB, Belgium

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED 482 082

Registar
Workshops
Sessions
Speakers
Interactions
Demonstrations
Exhibits
Events
Best of the Web
Key Dates
Seattle
Sponsors

A&MI
Archives & Museum Informatics
2008 Murray Ave
Suite D
Pittsburgh, PA
15217 USA
info@archimuse.com
www.archimuse.com

Search A&MI

Join our Mailing List

Privacy

Updated: 02/27/2001 11:24:54

Abstract

Museums have always been, sometimes directly and often indirectly, a key resource of arts and cultural heritage information for the classroom educator. The Web now offers an ideal way of taking this resource beyond the traditional textbook or school visit. While museums around the globe are embracing the web and putting virtual exhibitions, cultural databases and archives on-line, the educator (or user in general) is still facing the daunting task of integrating this material into an active document, course, curriculum or presentation. This paper reports on the construction of a personalized theme creation engine as a possible catalyst to the active use in secondary education in Europe of digital media published on-line by selected museums.

The HyperMuseum Theme Generator System (TGS) is part of the HyperMuseum system, a European virtual museum portal (Fig. 1). Its function is to assist in the creation of so-called *personalized themes*. A personalized theme intends to allow the end user to bring together a unique collection of multimedia objects from the HyperMuseum Server, and to create a personalized rendering of the perceived and/or recorded relationships between these objects, realized as a new multimedia document (website, PowerPoint® presentation, hypertext or Word® document etc.). The TGS supports this creative expression both during the *discovery* phase, exploring the collection and discovering thematic relationships, as well as the *realization* phase, the construction of the resulting documents. The primary target audience is non-expert users mainly from the secondary education community.

This work is partly supported by the Telematics Program (4th Framework) of the European Commission under Project nr. 3088 (HyperMuseum).

IR058768

BEST COPY AVAILABLE

Fig. 1: The Home Page of the HyperMuseum site http://www.HyperMuseum.com

A discovery system is grounded in a liberally linked *ontology service*. Ontologies, roughly speaking, are computer resources describing application domains in terms of standardized vocabularies, linking and categorizing those terms, for example as taxonomies. In our case, multiple linkage forms, selectable by the user, express and enhance the published collection of HyperMuseum digital media as a semantically linked network according to principles developed in the DOGMA project (http://starlab.vub.ac.be/dogma.htm) (Meersman, 1999a). Relationships between different objects are suggested and derived from the *metadata* that accompanies each object (as put there by subscribing museums), as well as previously published themes and general "background" ontologies. In exploring this network the user discovers a (new) path through the semantic links covering a theme, i.e. a novel set of relations between the objects. At any time he can export the newly discovered objects into the TGS creator tool, where the realization of the theme is ongoing. This creation phase concentrates on expressing and constructing the new relationships between the objects. It maintains a level of abstraction separate from the details of implementation in a specific medium, e.g. as an on-line Web presentation.

Overall HyperMuseum Functionality

The overall architecture of the complete HyperMuseum system consists conceptually of three different parts: The Museum Data Centers, The HyperMuseum Service Center (HSC) and the HyperMuseum Client (Fig. 2).

BEST COPY AVAILABLE

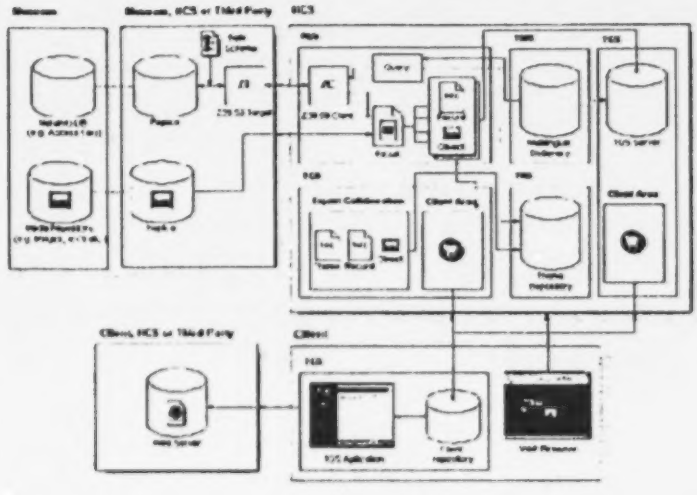

Fig. 2: Architectural overview of the Hypermuseum system.

The philosophy of the HyperMuseum is that requirements for participating museums should be minimal. It was felt that at this experimental stage the participating cultural institutions as well as most of the target market would not be inclined to dedicate the funds necessary for data conversion to a specific application. Therefore, existing data systems should if possible be reused without a requirement for conversion or standardization. Museum data can be replicated to the HSC servers, or be hosted by the museum itself on an accessible site. Media objects are served through a normal HTTP Web server, while corresponding data records are consulted through a Z39.50 server, with the HSC acting as a Z39.50 client (ANSI/NISO 1995). For each museum database, a Z39.50 profile is developed to map the existing (meta-)data fields onto a common structure. At the HSC side the retrieved data is converted into an XML format, and packaged in a ZIP file together with its corresponding digital media files. The result of this operation is referred to as an HMRecord file and will become the basic operational unit inside the rest of the HyperMuseum system.

The first service offered by the HSC is a traditional media consultation service called the Resource Discovery System (RDS) adapted from Aquarelle (Michard 1998). The RDS offers a Web interface to the virtual HyperMuseum collection, allowing the user to query for and retrieve information on the media objects. Several query methods are available, from generic free text search to structured search in specific contexts. This service is supported by a Terms Management System (TMS), offering assisted access to controlled vocabularies and multilingual dictionaries to facilitate intra linguistic retrieval.

The second service, the Theme Collaboration Service (TCS) is targeted towards the professional community. This part of the HyperMuseum eventually wants to offer a common workspace for creating themes based on the records in the HyperMuseum collection. This resulting theme files are once again based on an XML formalism, which allows the accompanying Theme Repository System (TRS) to make a collection of themes available to the public in a maintainable fashion. Both the Resource Discovery System and the Theme Repository

BEST COPY AVAILABLE

System are accessible with a Web browser at http://www.HyperMuseum.com (see also Fig. 1).

A third service, the Theme Generation Service, is the focus of this paper and will be explained in detail below

Rationale of the Theme Generation System

Experimenting with technology that can enhance the use of museum digital resources in secondary education is one of the main goals of the Hypermuseum project, and the central motivation underlying the creation of the TGS. We try to provide an alternative to the *passive* use of museum Web sites, which we might characterize as educational browsing, and support a more *active* type of usage, by letting the user explore and create with the objects of study.

Museum Web sites tend to fall into two categories. On the one hand there is the database approach to the holdings. The users are confronted with a search/query-like interface that allows them to get to the various assets in the digital collection. This approach has as an advantage in that it can easily offer access to the entire digital portfolio of the museum, and knowledgeable exploration can be made highly efficient because of the structured nature of the approach. Furthermore, the elements are more often presented without a contextual bias, allowing the users to appreciate the items from a personal background.

This approach is supported in the HyperMuseum through the RDS/TMS.

A second approach is that of the digital exhibition. Selected holdings are presented in a thematic context. This approach allows the *exhibitor* to convey a personal or institutional vision of a chosen set of media items. The educational and appreciative benefits of a professionally presented theme are evident and are supported in the HyperMuseum by the TCS/TRS.

With the Theme Generator System (TGS) we are experimenting with a third approach. From a previous project, Web For Schools (http://wfs.vub.ac.be) (Van Assche 1998) we had observed the need and enthusiasm from the secondary school educational community for uses of digital media in which there was a constructive component. This hands-on learning approach has become ingrained in the curricula of several European countries. Students not only observe, but also actively create and experiment with the educational materials on offer. The main goal of the TGS is to allow the discovery and the realization of a student's personal theme, based on the materials offered through the HyperMuseum, that can be used for a transient presentation to his peers in the classroom, and that can optionally be further elaborated into a more permanent project for sharing with others.

To support this approach, there must be access to the HyperMuseum resources that allows for the *discovery* of a theme. This is clearly different both from the strict catalogue approach and from the pre-packaged thematic approach.

The implementation of the Theme Generator is also intended as an early example of a practical *ontology-based*, or rather, in this case, *ontology-assisted* software tool. Tools of this kind assist the user in the creative

process by suggesting (or limiting) choices during the activities of query, search, design or composition of desired results by conceptually (semantically) linking linguistic elements of these activities (query text, figure captions, documentation, ...) with other documents and elements through "common" thesauri, lexicons, dictionaries, ... that cover the domain under consideration. Such linguistic resources may be quite general (i.e. domain-unspecific), as in the case of WordNet (Miller 1990), or alternatively quite specific to a particular domain. The latter kind of resource, now often referred to as a (*domain*) *ontology*, at present is much harder to come by. As an example of such an as yet hypothetical domain ontology, consider a (partial or comprehensive) listing of all the technical term linkages that may plausibly occur in the context of "restoration of medieval paintings". In fact, one of the primary goals of the DOGMA project (Meersman 1999b) at VUB STARLab (http://starlab.vub.ac.be) is to define a formalism, method and representation for such domain ontologies so they may be stored and integrated on an *ontology-server* and then retrieved, consulted and "plugged" into applications.

The main purpose - and advantage - of an ontology-based approach is to make the implementation of an application such as the TGS as independent of the data semantics (*meaning*) as is currently feasible, somewhat abstractly comparable in purpose to the way databases were invented to make applications maximally independent of data *structure*. While we certainly cannot claim that the TGS implementation presents a prototypical solution for this, we trust that it adequately illustrates the underlying principles - as well as offering an already useful tool that is extensible according to these principles, as discussed below in more detail.

Assisting systems with ontologies is not limited to the application under consideration. E.g. In the DOGMA research context also, library search systems as well as database (reverse-) engineering tools with ontological support are studied.

Supporting Personal Theme Discovery

While the database/catalogue approach to accessing museum record data is highly efficient in a *search* type of application, it is far from optimal in assisting in the discovery of the relations between the different elements. The thematic approach is strong in showing a possible linkage, but forces the user into considering a possibly very interesting set of relations in a pre-packaged format. For thematic discovery, we wanted an application, the Theme Hunter, which would support the user in fluidly discovering the patterns of possible, but as yet unrealized, themes.

We set out to construct a Web of relationships among the different objects in the HyperMuseum. In such a structure, the user could explore the space of relations by browsing this network from point to point, until a theme would emerge as an idea or pattern. At the same time, the user should be able to select the objects encountered for use in the realization of the theme.

For finding the possible relationships between the different objects in the HyperMuseum, we started from the textual metadata that is provided with each of the museum media objects (e.g. physical descriptions, materials, caption). We concentrated on the textual fields in the records.

Normalization of the words by using a stemming algorithm (Porter 1980) gives us a set of relations to words for each media-object. Bringing all the objects with their respective word links together in a graph creates a browsable structure (Fig.3), where one can move from object to object through the shared word links.

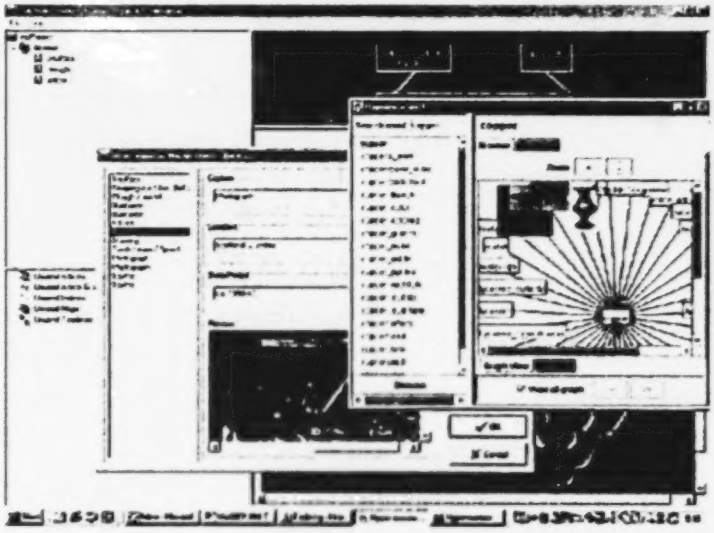

Fig. 3: Screenshot of the Theme Generator System User Interface showing the HMPPG (background) with the Theme Hunter (foreground)

While the browsing interface might be more adapted to the task of theme discovery, with only the word relations from the records we have done little but facilitate iterative searching of the record database. In order to get more depth in the possible theme exploration, we add in the system background ontological resources. These can provide for semantic links between the words that were linked to the records. Since this operation will provide the context for exploring thematic space, the choice of resources should be carefully matched with the intended audience and application. For our initial prototype we used an adaptation of the WordNet lexical database (Miller 1990).

The WordNet lexicon is centered around *synsets*, lists of synonyms that are intended to represent a meaning. Each synset is related to other synsets through semantic links. The types of links include among others *hyponym*, *hypernym*, *meronym*, *holonym* and *antonym* relations. Since a word can have more than one meaning, it can be part of more than one synset.

By extending our graph to include the word relations obtained from WordNet, we get a significantly richer linkage structure in our system. We no longer need to navigate through coinciding words, but can resolve synonyms and go into generalizations and specializations, opposites and hierarchies. This brings a qualitative change in the support of theme discovery. In the next section, we will delve deeper into the consequences of this type of facility.

A third source of relations between the objects and the graph can be

BEST COPY AVAILABLE

mined from *realized themes*. Since the HyperMuseum Theme Repository stores themes as XML documents, with embedded links to the media records, it is relatively straightforward to get the context of the reference and construct links in our graph between the context words and the objects. This expands the number of meaningful links (within this context) that exist between an object and the rest of the graph.

Supporting Personal Theme Generation

In general, one performs theme discovery with the purpose of communicating it. While it is certainly feasible to use a modern content creation package to construct multimedia presentations, there are tasks specific in our context, for example generating timelines or maps, that would clearly benefit from specialized automated support. This need has also been identified and reported in similar projects such as (Buchanan 1999). The Theme Generation System contains a client side application, the HyperMuseum Personal Project Generator (HMPPG), whose primary goal will be the facilitation of this aspect.

Since the theme itself should be kept independent of the details of the medium in which is it rendered, we tried to keep a separation between the conceptual creation of the theme (and in particular the ontology it uses), and its instantiation in a particular environment (e.g. as a Web site or Word document). While the current prototype only contains generators for Web site creation (as sets of HTML files), the implementation is such that the generation of different output types such as e.g. 3D virtual exhibitions as described in (Alonzo 2000) can be achieved through the inclusion of a new set of generators. The user constructs a theme based on conceptual objects such as groupings or indices that in one medium may be generated as Web pages and navigational menus, in another as virtual rooms and floor plans, and in yet another as chapters and a table of contents.

Realizing a theme based on museum objects is typically the creation of a new (thematic) dimension on the set of these objects. This new dimension does not replace the existing time or geographical dimension, but adds to these more objective references. Often the theme woven from a string of objects described in the thematic light will benefit from being also placeable and navigable along a timeline, or on a geographical map. Indeed, these factual data are often present in the HyperMuseum records, but the inclusion of these perspectives with general tools might be a laborious task. The HMPPG prototype generator automates the generation of timelines as HTML clickable maps, and a tool that automates placement of objects on geographical maps. These geographical maps can be created through a separately developed tool, GeoMap Editor, which is included in the HyperMuseum TGS toolset. Both these systems are based by default on the extraction of metadata from the HyperMuseum records. However, all project data can also be entered and edited manually for each instance, should this be necessary.

All elements in the HMPPG have themselves time and location *meta-data*. This makes it possible to cascade these elements into each other in essentially limitless combinations. One can easily construct timelines of maps and vice versa. Changing the order in the navigational hierarchy is a simple matter of dragging and dropping items from one place to another. The system is constantly checking for constraints (e.g. in order for an element to be placed on a timeline, it needs to have a date or

period associated), and will prompt for missing data as needed. This feature makes adding and changing different perspectives relatively straightforward.

At all times the user can observe what the element being constructed will look like with the currently selected generator. This provides immediate visual feedback. At any time more data can be fetched from the HyperMuseum through the Theme Hunter or the RDS, or new media files canbe added from the local file system. Most popular media types for images, audio and video are supported. The application can be used by itself to construct presentations. The output produced by the current prototype Web generator is standard HTML and can be further processed with other packages.

Architecture of the Theme Generation System

The Theme Generator System (Fig. 4) can be broken down into two major parts. One part is responsible for the theme realization and is referred to as the Hypermuseum Personal Project Generator. The second part is referred to as the Theme Hunter, and is responsible for theme discovery support. While both are very different in their implementation details, they integrate seamlessly from the user's point of view, with no noticeable transitions.

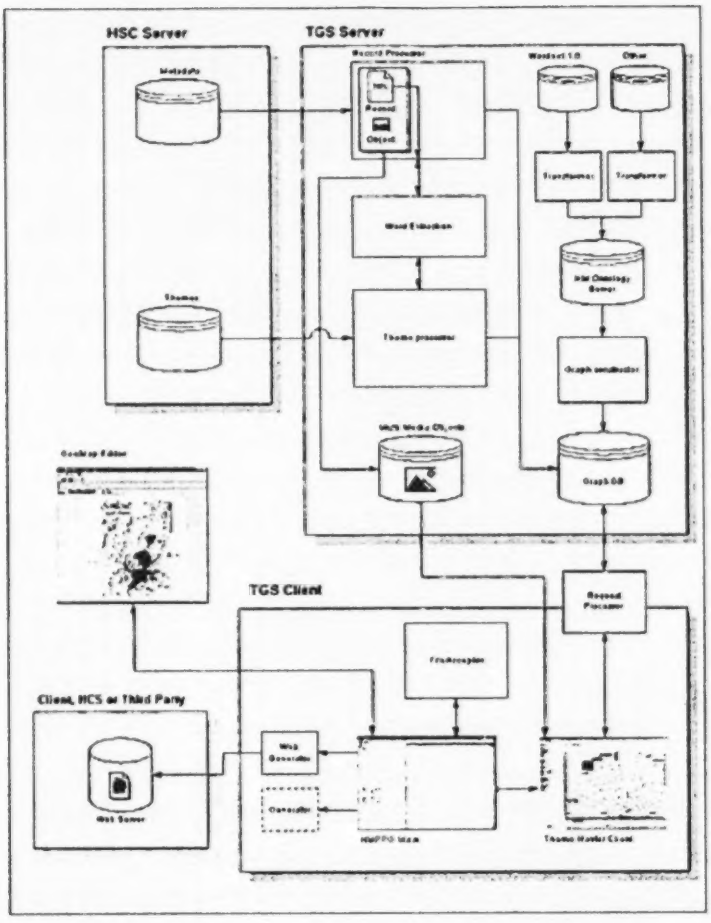

Fig. 4: Overview of the Theme Generator System Architecture

The Theme Hunter is a classical client-server application. The centerpiece of this system is a relational database system which houses the object-ontology graph described above. Client applications are either delivering new graph elements, or querying the graph. For practical purposes, the delivery of the actual multimedia files is handled through a standard Web server. Since the only point of interface is the graph in the database and the media files on the webserver, any application that can deliver graph extensions can be added to the system. This feature, together with the localization of format-dependent code in the HMPPG, will make it possible for the system to easily migrate to or be extended with new data formats.

As seen in the database diagram(Fig 5), the graph is represented as a simple set of binary relations between entities. The entities are typed to distinguish e.g. between words and object references, and associated with a source. This source is itself typed, and in our current system there are three types: the Hypermuseum Records, the Hypermuseum Themes and WordNet. Each source type has a set of relations that can be further classified into relation types for easy property attribution. From the Theme Hunter client, the advanced user can configure the system to work only on certain sets of sources, or take into account only certain

BEST COPY AVAILABLE

relations.

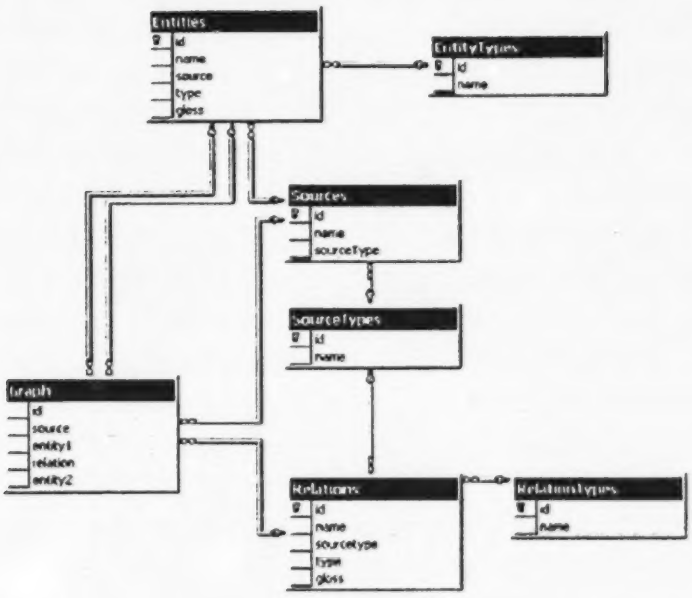

Fig 5.: Theme Generator System Graph Database Diagram

On the server side there are three main data delivery modules, each responsible for extracting information from the different data sources: the Hypermuseum Records, the Hypermuseum Themes, and the WordNet 1.6 data files. For the WordNet intaker, we start from the DOGMA representation of WordNet 1.6, which is in content identical to the WordNet data files, but allows for easier manipulation. From this representation we can build the required graph representation through direct SQL manipulations. We build a semantic net graph centered around the words, as opposed to the synsets. We ambiguate the meaning structure by flattening the word-meaning relations into a single node, and transform the synset objects into explicit synonym relationships between the member words.

For the Themes and the Records, the first step is extracting the field data that is to be considered for word mining. This operation is made easier by the fact that both use an XML representation, making the parsing and validation of the files more convenient. From this step the relation between the object and the possible word set is determined. Text data is passed through the word extraction process, in the current prototype a chunking (cutting the sentence into pieces) and stemming operation. There are many opportunities here for the application of advanced Natural Language Processing techniques that could improve both the accuracy and the depth of the text mining operation. The last stage in our current mining pipeline matches the words against WordNet 1.6 to prune the false results out of the previous pipeline stages. The resulting words are then inserted into the graph, together with their respective relations to the objects. Items recovered form Theme files are only included if the record they refer to was previously processed by the system.

BEST COPY AVAILABLE

The part of the Theme Hunter that is visible to the end user is conceived as a 3-tier application on the graph database. The front end is a GUI based application (Fig. 3) where the user can search for terms to start exploring in the graph, and graphically browse the graph by simple mouse clicks. When interesting objects are discovered, they can be downloaded for processing with the HMPPG by a simple right click on the image. The Theme Hunter and the HMPPG work completely asynchronously, so realization and discovery can be fully intermixed in the creative process.

Behind the scenes the first tier constructs a series of request objects that are passed on to the middle tier. The middle tier interprets these requests and builds database queries out of these. With these queries it extracts the information out of the graph database, and the result is transformed into a reply object and passed back to the first tier, where the result is unpacked and displayed. The protocol governing this interaction is stateless, ensuring that the process is robust to client or server failure and economically scalable on the server side. Furthermore, both the request and reply objects that govern the first-middle tier interaction are serializable, so that the actual physical location of the middle tier is easily changeable. In our current setup we placed the middle tier on the client side, but it can with relatively minor efforts be moved to the server side should the need arise.

Our Theme Hunter implementation uses the Microsoft SQL Server v7.0, and has also been tested with the free Microsoft Data Engine 1.0 for cheaper deployment scenarios. Since the database operations are handled through SQL over a standard ODBC interface, there should be no problem to substitute for these choices. The database is complemented by a standard Web server, in our system Internet Information Server 4.0, for the delivery of the actual multimedia files to the clients. All code for this part of the system was written in Java 2 except for some of the DB conversion code that was written directly in SQL.

Like the Theme Hunter, the Hypermuseum Personal Project Generator also tries to isolate the dependencies on data formats. The HMPPG is a client side program that is constructed as an output-format-agnostic generation framework. Data input into the system can be from Hypermuseum record files that can originate either from interaction with the RDS or from the Theme Hunter, or from raw multimedia data supplied by the user from other sources. Due to its intense local data interaction, this application was developed mainly using the Borland C++ Builder system, since the current instantiations of the Java platform did not deliver the required performance at this point. (The Theme Hunter, being a network-data bound application, is not critically hindered by this). The dependency on the Hypermuseum record format is shielded by locating the parsing and extracting code into a separate program, the FileAcceptor. The responsibility of the FileAcceptor is to extract from a set of Hypermuseum records the relevant metadata and the media files, and produce these into an internal format on which the rest of the HMPPG will rely. (Since there is no user interface involved here, we could develop this program in Java 2).

To achieve data format independence on the generation side, two requirements had to be fulfilled.

First, all the operations on the data, apart from the generated

representations, had to be implemented as operations on the conceptual structures. Second, the generators could not directly be linked into the rest of the program. Instead, a COM (Rogerson 1997) interface was defined that specifies which services a generator has to provide to the HMPPG. Since the HMPPG acts towards the generators as a COM client, each generator object implementing this interface can be added to the one delivered with the current prototype.

The current HMPPG generators can work with most of the standard formats for audio, video and imagery. Every format supported by the Windows Media Player and the Internet Explorer controls may be used in the prototype. The GeoMap file format for geographical data is supported through the GeoMap Editor tool that also acts as a COM server towards the HMPPG for handling GeoMap operations.

We opted for this defensive position with regards to changing data and output format requirements because of the current absence of stable and dominant standards for this field. Since this situation is currently receiving quite some interest, we hope to be able to incorporate the emerging standards into the system in the future.

Discussion and Results of Initial Experimentation

As stated above, the choice of background ontological resources is critical to the intended use of the system. The choice of the WordNet lexicon was actually one of convenience and availability, but first reactions to the prototype indicate that the choice may have unexpected benefits.

Since WordNet has attempted to model the lexical knowledge of a native speaker of English, its vocabulary is quite broad but non-specific. As a result, the user is generally not confronted with jargon which might have been more accurate or meaningful to the professional user, but could have alienated or derailed our untrained user in the process of making free associations. Secondly, while our present crude word harvesters can extract the words, they cannot disambiguate them to their intended meaning (e.g. it can not distinguish between the use of bank as the place that takes your money as opposed to the bank at the side of the river). Rather than an expected disadvantage, this "feature" in the context of our application surprisingly seems to bepositive, since these mismatches seem to be one of the best sources of theme "triggers". Even the fact that our current main view in the browser (there is a graphical main view and a textual view) does not show the relation types has triggered surprise, reflection and discovery in the first test when users unexpectedly came across links they at first thought were not supposed to be there (e.g. the link between "rich" and "tasteful"). For some instances at least, less could indeed be more when selecting background ontologies.

The previous point is not intended to be a dogmatic principle. We could point to the lack of agentive links (e.g. there is no relation between "baker" and "bread") as an obvious shortcoming of our current system.

Once the background ontology is in place and the graph generated, the choice of how much of it to deploy can only be discovered through experimentation. In our current Theme Hunter, the advanced user can include or exclude both types of relations and sources. Even so, depending on the scenario we can foresee a need to further reduce the

amount of presented information. We are currently investigating the introduction of additional retrieval constraints such as a maximal distance rule (words are only included in the graph within at least few "hops" of a media item) to bring more focus on finding themes that will be "realizable" with the HyperMuseum data at hand. These criteria are dependent on the specifics of the usage scenario and will have to be evaluated on a case-by-case basis.

As far as the records themselves are concerned, the HyperMuseum approach of not requiring rigorous data standardization undoubtedly has the advantage of lowering the threshold for museum participation. The flip side of this coin is, of course, that since there are not yet standard representational formats for things like dates or geographic locations, the correct interpretation of these becomes difficult (time) to very difficult (geographic). A common reference to standardized vocabularies such as the Getty vocabularies (Lanzi 1998) could augment the quality of information extraction by the different tools.

Conclusion

The HyperMuseum Theme Generator System aims to be a contribution to the exploration of new uses of museum online digital information in, among other environments, the secondary education curricula in Europe. By allowing a novel way of exploring collections, supporting both the discovery of new thematic dimensions by the non-professional as well as the rapid creation of the presentation of the theme, we hope to address concrete needs of the intended audience.

By localizing the dependencies on specific formats, both on the data delivery side and the document generation side, we hope to have provided a technical platform that can be flexibly adapted for usage in many environments.

Acknowledgements

This work is supported by the European Commission's EU Telematics Application Project HYPERMUSEUM (Project nr. 3088). We would also like to acknowledge contributions by our museum partners in this project, the Oxfordshire County Council museums (Oxford, UK), the Musée Calvet (Avignon, FR) and the Galleria Degli Uffizi (Firenze, I) . In particular the dataset developed by the National Museum of Scotland (Edinburgh, UK) provided an excellent ontology testing tool.

References

Alonzo, F., Garzotto, F., and Valenti, S. (2000) 3-D Temporal Navigation on the Web: How to Explore a Virtual City along Multiple Historical Perspectives., In D. Bearman & J. Trant (Eds.) *Museums and the Web 2000 Proceedings.*, Archives & Museum Informatics, 2000.

ANSI/NISO (1995) *Information Retrieval: Application Service Definition and Protocol Specification.* Available at http://lcweb.loc.gov/z3950/agency/document.html

Buchanan, S. (1999) Unlocking the Treasure Chest Using SCRAN Tools. In D.Bearman & J. Trant (Eds.) Museums and the web 1999

Proceedings., Archives & Museum Informatics, 1999.

Lanzi, E. (1998) *Introduction to Vocabularies, Enhancing Access to Cultural Heritage Information*. Los Angeles: Getty Trust Publications.

Meersman, R. (1999a) Semantic Ontology Tools in Information Systems Design. In Z. Ras & M. Zemankova (Eds.) *Proceedings of the ISMIS'99 Conference*. Heidelberg: Springer Verlag.

Meersman, R. (1999b). The use of lexicons and other computer-linguistic tools in semantics, design and cooperation of database systems. In: Yanchun Zhang, M. Rusinkiewicz, Y. Kambayashi (eds.), *Proceedings of the Conference on Cooperative Database Systems (CODAS'99)*. Heidelberg: Springer Verlag.

Michard A., Christophides V., Scholl M., Stapleton M., Sutcliffe D., and Vercoustre A-M. (1998) The Aquarelle Resource Discovery System., *Journal of Computer Networks and ISDN Systems*, 30(13):1185--1200.

Miller, G. A., Beckwith R., Felbaum C., Gross D., and Miller K., (1990). Introduction to WordNet : An On-line Lexical Database., *International Journal of Lexicography*, 3, (4), 235 - 244.

Porter, M.(1980) An algorithm for suffix stripping. *Program*, 14, (3), 130-137.

Rogerson, D. (1997) Inside COM. Redmond: Microsoft Press.

Van Assche, F (1998) (Ed.). *Using the WWW in Secondary Schools*. Leuven: Acco.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

[X] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

[ ] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").